

# Low Treewidth Embeddings of Planar and Minor-Free Metrics

Arnold Filtser  
Bar-Ilan University  
arnold.filtser@biu.ac.il

Hung Le  
University of Massachusetts at Amherst  
hungle@cs.umass.edu

**Abstract**—Cohen-Addad, Filtser, Klein and Le [FOCS’20] constructed a stochastic embedding of minor-free graphs of diameter  $D$  into graphs of treewidth  $O_\epsilon(\log n)$  with expected additive distortion  $+\epsilon D$ . Cohen-Addad et al. then used the embedding to design the first quasi-polynomial time approximation scheme (QPTAS) for the capacitated vehicle routing problem. Filtser and Le [STOC’21] used the embedding (in a different way) to design a QPTAS for the metric Baker’s problems in minor-free graphs. In this work, we devise a new embedding technique to improve the treewidth bound of Cohen-Addad et al. exponentially to  $O_\epsilon(\log \log n)^2$ . As a corollary, we obtain the first efficient PTAS for the capacitated vehicle routing problem in minor-free graphs. We also significantly improve the running time of the QPTAS for the metric Baker’s problems in minor-free graphs from  $n^{O_\epsilon(\log(n))}$  to  $n^{O_\epsilon(\log \log(n))^3}$ .

Applying our embedding technique to planar graphs, we obtain a deterministic embedding of planar graphs of diameter  $D$  into graphs of treewidth  $O((\log \log n)^2/\epsilon)$  and additive distortion  $+\epsilon D$  that can be constructed in nearly linear time. Important corollaries of our result include a bicriteria PTAS for metric Baker’s problems and a PTAS for the vehicle routing problem with bounded capacity in planar graphs, both run in *almost-linear* time. The running time of our algorithms is significantly better than previous algorithms that require quadratic time.

A key idea in our embedding is the construction of an (exact) emulator for tree metrics with treewidth  $O(\log \log n)$  and hop-diameter  $O(\log \log n)$ . This result may be of independent interest.

**Index Terms**—metric embedding, minor-free graphs, planar graphs, vehicle routing problem, PTAS

## I. INTRODUCTION

Metric embedding is an influential algorithmic tool that has been applied to many different settings, for example, approximation/sublinear/online/distributed algorithms [6], [27], [69], [73], machine learning [57], computational biology [62], and computer vision [9]. The fundamental idea of metric embedding in solving an algorithmic problem is to embed an input metric space to a host metric space that is “simpler” than the input metric space, solve the problem in the (simple) host metric space and then map the solution back to a solution of the input metric space. In this algorithmic pipeline, the structure of the host metric space plays a decisive role.

In their seminal result, Fakcharoenphol, Rao and Talwar [44] (improving over Bartal [11], [12], see also [13])

constructed a stochastic embedding of an arbitrary  $n$ -point metric space to a tree with expected multiplicative distortion  $O(\log n)$ ; the distortion was shown to be optimal [11]. One may hope to get better distortion by constraining the structure of the input metric space, or enriching the host space. Shattering such hopes, Carroll and Goel [28] (implicitly) showed that there in an infinite family of planar graphs, such that every deterministic embedding into treewidth  $n^{\frac{1}{3}}$  graphs will have multiplicative distortion  $\Omega(n^{\frac{1}{3}})$ . Furthermore, Chakrabarti, Jaffe, Lee, and Vincent [29] showed that any stochastic embedding of planar graphs into graphs with *constant* treewidth requires expected distortion  $\Omega(\log n)$ . Carroll and Goel [28] showed a more general trade-off: any stochastic embedding of planar graphs with expected distortion  $c \geq 1$  requires treewidth  $\Omega(\log(n)/c)$ . In particular, any stochastic embedding with expected distortion  $\sqrt{\log n}$  has treewidth  $O(\sqrt{\log n})$ .

Bypassing this roadblock, Fox-Epstein, Klein, and Schild [56] studied *additive embeddings*: a  $\Delta$ -additive embedding  $f : V(G) \rightarrow V(H)$  of a graph  $G$  to a graph  $H$  is an embedding such that for every  $u, v \in V(G)$ :

$$d_G(u, v) \leq d_H(f(u), f(v)) \leq d_G(u, v) + \Delta .$$

The parameter  $\Delta$  is the *additive distortion* of the  $\Delta$ -additive embedding  $f$ . Fox-Epstein et al. [56] showed that planar graphs of diameter  $D$  admit a (deterministic)  $(\epsilon D)$ -additive embedding into graphs of treewidth  $O(\epsilon^{-c})$  for some universal constant  $c$ . The constant treewidth bound (for a constant  $\epsilon$ ) sharply contrasts additive embeddings with multiplicative embeddings (where the treewidth is polynomial). Their motivation for developing the additive embedding was to design PTASes for the metric Baker’s problems in planar graphs.

In a seminal paper [10], Baker designed PTASes for several problems in planar graphs such as independent set, dominating set, and vertex cover, where vertices have non-negative measures (or weights). Note that these problems are APX-hard in general graphs. Baker’s results subsequently inspired the development of powerful algorithmic frameworks for planar graphs, such as deletion decomposition [10], [38], [43], contraction decomposition [39], [41], [70], and bidimensionality [37], [53]. Metric Baker’s problems generalize Baker’s problems in that vertices in the solution must be at least/at most a distance  $\rho$  from each other for some input parameter  $\rho$ . The most well-studied examples of metric Baker’s problems include  $\rho$ -independent set,  $\rho$ -dominating set,  $(k, r)$ -center.

The first author was supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1042/22). The second author was supported by the National Science Foundation under Grant No. CCF-2121952.

Metric Baker problems have been studied in the context of parameterized complexity [24], [40], [66], [74] where (a) the input graphs are restricted to subclasses of minor-free graphs, such as planar and bounded treewidth graphs, and (b) parameters such as  $\rho$  and/or the size of the optimal solution are small. When  $\rho$  is a constant, Baker’s layering technique can be applied to obtain a linear time PTAS for unweighted planar graphs [10] and efficient PTASes for unweighted minor-free graphs [38]. However, the most challenging case is when the graph is weighted, and  $\rho$  is part of the input; even when restricted to bounded treewidth graph, (single-criteria) PTASes are not known for metric Baker problems. Marx and Pilipczuk [74] showed that, under Exponential Time Hypothesis (ETH),  $\rho$ -independent/dominating set problems cannot be solved in time  $f(k)n^{o(\sqrt{k})}$  when the solution size is at most  $k$ . As observed by Fox-Epstein et al. [56], the result of Marx and Pilipczuk [74] implies that, under ETH, there is no (single-criteria) efficient PTAS for  $\rho$ -independent/dominating set problems in planar graphs. However, for the case of uniform measure (i.e.  $\forall v, \mu(v) = 1$ ), a (non-efficient) PTAS can be obtained via local search [50].

Fox-Epstein et al. [56] bypassed the ETH lower bound by designing a *bi-criteria* efficient PTAS for  $\rho$ -independent/dominating set problems in planar graphs using their additive distortion embedding of planar graphs into bounded treewidth graphs. Since the treewidth of their embedding is  $O(\epsilon^{-c})$  for some constant  $c \geq 19$  (see Section 6.5 in [56]), and the running time to construct the embedding is  $n^{O(1)}$  for an unspecified constant in the exponent due to the embedding step, the running time of their PTAS is  $2^{O(\epsilon^{-(c+1)})}n^{O(1)}$  [56]. In their paper, they noted:

“Admittedly, in our current proof, the treewidth is bounded by a polynomial of very high degree in  $1/\epsilon$ . There is some irony in the fact that our approach to achieving an efficient PTAS yields an algorithm that is inefficient in the constant dependence on  $\epsilon$ .”

Given the state of affairs, the following problem arises:

**Question 1.** *Can we design a PTAS for metric Baker’s problems with (almost) linear running time? Can we obtain a PTAS with a more practical dependency on  $\epsilon$ ?*

Cohen-Addad, Filtser, Klein, and Le [34] studied additive embeddings in a more general setting of  $K_r$ -minor-free graphs for any fixed  $r$ . They proved a strong lower bound on treewidth against *deterministic* additive embeddings. Specifically, they showed (Theorem 3 in [34]) that there is an  $n$ -vertex  $K_6$ -minor-free graph such that any deterministic additive embedding into a graph of treewidth  $o(\sqrt{n})$  must incur a distortion at least  $\frac{D}{20}$ . On the other hand, they showed that randomness helps reduce the treewidth exponentially. In particular, they constructed a *stochastic* additive embedding of  $K_r$ -minor-free graphs into graphs with treewidth  $O(\frac{\log n}{\epsilon^2})$  and *expected* additive distortion  $+\epsilon D$ . Specifically, there is distribution  $\mathcal{D}$  over dominating embeddings (i.e. no distances shrink) into treewidth  $O(\frac{\log n}{\epsilon^2})$  graphs such that

$$\forall u, v, \mathbb{E}_{f, H_f}[d_{H_f}(f(u), d(v))] \leq d_G(u, v) + \epsilon D.$$

Their primary motivation was to design a quasi-polynomial time approximation scheme (QPTAS) for the bounded-capacity vehicle routing problem (VRP) in  $K_r$ -minor-free graphs. In this problem, we are given an edge-weighted graph  $G = (V, E, w)$ , a set of clients  $K \subseteq V$ , a depot  $r \in V$ , and the capacity  $Q \in \mathbb{Z}_+$  of the vehicle. We are tasked with finding a collection of tours  $\mathcal{S} = \{R_1, R_2, \dots\}$  of *minimum cost* such that each tour, starting from  $r$  and ending at  $r$ , visits at most  $Q$  clients and every client is visited by at least one tour; the cost of  $\mathcal{S}$  is the total weight of all tours in  $\mathcal{S}$ . The VRP was introduced by Dantzig and Ramser [35] and has been extensively studied since then; see the survey by Fisher [52]. The problem is APX-hard, as it is a generalization of the Travelling Salesperson Problem (TSP) when  $|Q| = |V|$ , which is APX-hard [80], and admits a constant factor approximation [61]. To get a  $(1 + \epsilon)$ -approximation, it is necessary to restrict the structures of the input graph. Fundamental graph structures that have long been studied are low dimensional Euclidean (or doubling) spaces, planarity and minor-freeness.

When the capacity  $Q$  is a part of the input, a QPTAS for VRP in Euclidean space of constant dimension is known [4], [36]; it remains a major open problem to design a PTAS for VRP even for the Euclidean plane. For trees, a PTAS was only obtained by a recent work of Mathieu and Zhou [75], improving upon the QPTAS of Jayaprakash and Salavatipour [65]. No PTAS is known beyond trees, such as planar graphs or bounded treewidth graphs. For the *unsplittable* demand version of the problem on trees, Becker [17] showed that the problem is APX-hard.

Going beyond trees, it is natural to restrict the problem further by considering *constant*  $Q$ . In this regime, Becker et al. [19] designed the first (randomized) PTAS for bounded-capacity VRP in planar graphs with running time  $n^{O_\epsilon(1)}$ , improving upon the earlier QPTAS by Becker, Klein and Saulpic [18]. Recently, Cohen-Addad et al. [34] obtained the first efficient PTAS for the problem in planar graphs with running time  $O_\epsilon(1)n^{O(1)}$ . Their algorithm uses the embedding of [56] as a blackbox, and hence, suffers the drawback of the embedding: the exponent of  $n$  in the running time is unspecified due to the embedding step.

In  $K_r$ -minor-free graphs, Cohen-Addad et al. [34] designed a QPTAS for bounded-capacity VRP in  $K_r$ -minor-free graphs using their (stochastic) embedding of  $K_r$ -minor-free graphs into graphs with treewidth  $O(\frac{\log n}{\epsilon^2})$ . More precisely, their algorithm has running time  $(\log n)^{\text{tw}}n^{O(1)}$  where  $\text{tw}$  is the treewidth of the embedding. Thus, any significant improvement over treewidth bound  $O(\frac{\log n}{\epsilon^2})$ , such as  $O(\log n / \log \log n)$ , would lead to a PTAS. Their work left the following questions as open problems.

**Question 2.** *Can we devise an additive embedding of  $K_r$ -minor-free graphs into graphs with treewidth  $O(\log(n) / \log \log(n))$ ? Can we design a PTAS for the bounded-capacity VRP in  $K_r$ -minor-free graphs? Can we improve the running time of the PTAS for the bounded-*

capacity VRP in planar graphs to (almost) linear?

We remark that the dependency on  $\epsilon$  of all known approximation schemes for bounded-capacity VRP problem is *doubly exponential* in  $1/\epsilon$  [19], [34]. Reducing the dependency to singly exponential in  $1/\epsilon$  is a fascinating open problem.

One drawback of the stochastic embedding with additive distortion is that it cannot be applied directly to design (bicriteria) PTASes for metric Baker problems in minor-free graphs. To remedy this drawback, the authors [50] recently introduced *clan embedding* and *Ramsey type embedding* with additive distortions. A clan embedding of a graph  $G$  to a graph  $H$  is pair of maps  $(f, \chi)$  where  $f$  is a *one-to-many* embedding  $f : V(G) \rightarrow 2^{V(H)}$  that maps each vertex  $x \in V(G)$  to a subset of vertices  $f(x) \subseteq V(H)$  in  $H$ , called *copies* of  $x$  (where each set  $f(x) \neq \emptyset$  is not empty, and every two sets of  $x \neq y$  are  $f(x) \cap f(y) = \emptyset$  are disjoint), and  $\chi : V(H) \rightarrow V(H)$  maps each vertex  $x$  to a vertex  $\chi(x) \in f(x)$ , called the *chief* of  $x$ . Furthermore,  $f$  must be *dominating*:

$$\forall x, y \in V(G), \quad d_G(x, y) \leq \min_{x' \in f(x), y' \in f(y)} d_H(x', y').$$

A clan embedding  $(f, \chi)$  has additive distortion  $+\Delta$  if for every  $x, y \in V(G)$ :

$$\min_{x' \in f(x)} d_H(x', \chi(y)) \leq d_G(x, y) + \Delta \quad (1)$$

That is, there is some vertex in the clan of  $x$  which is close to the chief of  $y$ . Note that the distortion guarantee is in the worst case. That is, Equation (1) holds for every pair  $x, y \in V(G)$ , and every embedding  $f \in \text{support}(\mathcal{D})$  in the support.

A Ramsey type embedding is a (stochastic) one-to-one embedding in which there is a subset of vertices  $M \subseteq V$  such that every vertex is included in  $M$  with a probability at least  $1 - \delta$ , for a given parameter  $\delta \in (0, 1)$ , and for every vertex  $u \in M$ , the additive distortion of the distance from  $u$  to every other vertex in  $V$  is  $+\Delta$ .

The authors [50] showed previously that for any given  $K_r$ -minor-free graph and parameters  $\epsilon, \delta \in (0, 1)$ , one can construct a distribution  $\mathcal{D}$  over clan embeddings into  $O_r(\frac{\log^2 n}{\epsilon \delta})$ -treewidth graphs with additive distortion  $+\epsilon D$  ( $D$  being the diameter) and such that the expected clan size  $\mathbb{E}[|f(x)|]$  of every  $x \in V(G)$  is bounded by  $1 + \delta$ . For Ramsey type embeddings, the treewidth is also  $O_r(\frac{\log^2 n}{\epsilon \delta})$  for an additive distortion  $+\epsilon D$  [50].

Using the clan embedding and Ramsey type embedding, the authors obtained a QPTAS for metric Baker's problem in  $K_r$ -minor-free graphs with running time  $n^{O_r(\epsilon^{-2} \log(n) \log \log(n))}$  [50]. The precise running time of the algorithm is  $(\log n)^{O(\text{tw})} n^{O(1)}$  where  $\text{tw}$  is the treewidth of the embeddings. The key questions are:

**Question 3.** *Can we improve the treewidth bound  $O(\log^2(n))$  of the clan embedding and Ramsey embedding? Can we design a PTAS for metric Baker's problems in  $K_r$ -minor-free graphs?*

## A. Our Results

We provide affirmative answers to Question 1 and Question 2, while making significant progress toward Question 3.

We construct a new stochastic additive embedding of  $K_r$ -minor-free graphs into graphs with treewidth  $O_r(\frac{(\log \log n)^2}{\epsilon^2})$ . Our treewidth bound improves *exponentially* over the treewidth bound of Cohen-Addad et al. [34]. See Table I for a summary of new and previous embeddings.

**Theorem 1** (Embedding Minor-free Graphs to Low Treewidth Graphs). *Given an  $n$ -vertex  $K_r$ -minor-free graph  $G(V, E, w)$  of diameter  $D$ , we can construct in polynomial time a stochastic additive embedding  $f : V(G) \rightarrow H$  into a distribution over graphs  $H$  of treewidth at most  $O_r(\epsilon^{-2}(\log \log n)^2)$  and expected additive distortion  $\epsilon D$ .*

The main bottleneck for not having an almost-linear time algorithm for the embedding in minor-free graphs is that the best algorithm for computing Robertson-Seymour decomposition takes quadratic time [68].

Using our stochastic additive embedding, we design the first PTAS for the bounded-capacity vehicle routing problem in  $K_r$ -minor-free graphs; our PTAS indeed is efficient. This resolves Question 2 (more on the planar case later in Theorem 6).

**Theorem 2** (PTAS for Bounded-Capacity VRP in Minor-free Graphs). *There is a randomized polynomial time approximation scheme for the bounded-capacity VRP in  $K_r$ -minor-free graphs that runs in  $O_{\epsilon, r}(1) \cdot n^{O(1)}$  time.*

Our proof of Theorem 2 follows the embedding framework of Cohen-Addad et al. [34]. In a nutshell, they showed that if planar graphs with *one vortex* have an additive embedding with treewidth  $k(n, \epsilon)$ , then  $K_r$ -minor-free graphs have a stochastic additive embedding with treewidth roughly  $O_r(k(n, O_r(\epsilon^2)) + \log(n))$ . That is, the reduction incurs an additive factor of  $O(\log n)$ . Thus, there are two issues one has to resolve to reduce the treewidth to  $o(\log(n))$ : (i) construct an embedding of planar graphs with one vortex that has treewidth  $k(n, \epsilon) = o(\log(n))$  and (ii) remove the loss  $O(\log(n))$  in the reduction of Cohen-Addad et al. [34]. By a relatively simple idea, we could remove the additive term  $O(\log(n))$  in the reduction. We are left with constructing an embedding of planar graphs with one vortex, which is the main barrier we overcome here.

One can intuitively think of planar graphs with one vortex as *noisy* planar graphs, where the vortex is a kind of low-complexity<sup>1</sup> noise that affects the planar embedding in a local area, i.e., a face. From this point of view, we need an embedding of planar graphs that is *robust* to the noise. The embedding for planar graphs of Fox-Epstein et al. [56] relies heavily on planarity to perform topological operations such as cutting paths open, and decomposing the graphs into so-called *bars and cages*. As a result, adding a vortex to a planar graph makes their embedding inapplicable. Cohen-Addad et al. [34] instead use *balanced shortest path separators* in their embedding of planar graphs with one vortex. Using shortest path

<sup>1</sup>Low-complexity means that the vortex has bounded pathwidth.

TABLE I: Summary of current and previous embeddings into low treewidth graphs with additive distortion  $+\epsilon D$ . For planar graphs, we get an embedding whose treewidth has a minor dependency on  $n$ ; however our running time and dependency on  $\epsilon$  is much improved. For stochastic embeddings of  $K_r$ -minor-free graphs, we obtain an exponential improvement. For Ramsey-type and clan embeddings, we obtain a quadratic improvement. All known results for planar graphs, including ours, can be extended to bounded genus graphs with only a constant loss in the treewidth using the cutting technique described in [34].

Family	Type	Treewidth	Runtime	Ref
Planar	Deterministic	$O_r(\epsilon^{-1} \cdot \log n)$	$O_\epsilon(n^{O(1)})$	[42]
		$O_r(\epsilon^{-c}, c \geq 19)$	$O_\epsilon(n^{O(1)})$	[56]
		$O_r(\epsilon^{-1} \cdot (\log \log n)^2)$	$\epsilon^{-2} \cdot O(n)$	Theorem 3
$K_r$ -minor free	Stochastic	$O_r(\epsilon^{-2} \cdot \log n)$	$O_{\epsilon,r}(n^{O(1)})$	[34]
		$O_r(\epsilon^{-2} \cdot (\log \log n)^2)$	$O_{\epsilon,r}(n^{O(1)})$	Theorem 1
	Ramsey type / Clan	$(\epsilon\delta)^{-1} \cdot O_r(\log^2 n)$	$O_{\epsilon,r}(n^{O(1)})$	[50]
		$(\epsilon\delta)^{-1} \cdot O_r(\log n)$	$O_{\epsilon,r}(n^{O(1)})$	Theorems 7 and 8

separators results in an embedding technique that is robust to the noise caused by the vortex. However, their embedding has treewidth  $O(\log(n)/\epsilon)$  due to that recursively decomposing the graphs using balanced separators gives a decomposition tree of depth  $O(\log n)$ . This is a universal phenomenon in almost all techniques that rely on balanced separators: the depth  $O(\log n)$  factors in many known algorithms in planar graphs [8], [33], [42], [55], [58]. It seems that to get a treewidth  $o(\log n)$ , one needs to avoid using balanced separators in the construction.

Surprisingly perhaps, we can still use balanced shortest path separators to get an embedding with treewidth  $O((\log \log n)^2/\epsilon)$  for planar graphs with one vortex. Our key idea is to “shortcut” the decomposition tree of depth  $O(\log n)$  by adding edges between nodes in such a way that, for every two nodes in the tree, there is a path in the shortcut tree with  $O(\log \log n)$  edges, a.k.a.  $O(\log \log n)$  hops. To keep the treewidth of the embedding small, we require that the resulting treewidth of the decomposition tree (after adding the shortcuts) remains small; the decomposition tree is a tree and hence has treewidth 1. We show that we can add shortcuts in such a way that the resulting treewidth is  $O(\log \log(n))$ . The two factors of  $O(\log \log(n))$  — one from the hop length and one from the treewidth blowup due to shortcutting — result in treewidth of  $O((\log \log n)^2/\epsilon)$  of the embedding. We formulate these ideas in terms of constructing an emulator for trees with treewidth  $O(\log \log n)$  and hop-diameter  $O(\log \log n)$ ; see Section I-B for a more formal discussion. The main conceptual message of our technique is that it is possible to get around the depth barrier  $O(\log n)$  of balanced separators by shortcutting the decomposition tree. We believe that this idea would find further use in designing planar graph algorithms.

Applying our technique to planar graphs, we obtain an additive embedding with treewidth  $O((\log \log n)^2/\epsilon)$  that can be constructed in nearly linear time.

**Theorem 3** (Embedding Planar Graphs to Bounded Treewidth Graphs). *Given an  $n$ -vertex planar graph  $G(V, E, w)$  of diameter  $D$  and a parameter  $\epsilon \in (0, 1)$ , there is a deterministic embedding  $f : V(G) \rightarrow H$  into a graph  $H$  of treewidth  $O(\epsilon^{-1}(\log \log n)^2)$  and additive distortion  $+\epsilon D$ . Furthermore,  $f$  can be constructed in  $O(n \cdot \frac{\log^3 n}{\epsilon^2})$  time.*

Our embedding, while having a minor dependency on  $n$ , offers three advantages over the embedding of [56]. First, our embedding can be constructed in nearly linear running time, removing the running time bottleneck of algorithms that uses the embedding of [56]. Second, our embedding algorithm is much simpler than that of [56] as it only uses the recursive shortest path separator decomposition. Third, the dependency on  $\epsilon$  of the treewidth is linear, which we show to be optimal by the following theorem.

**Theorem 4** (Embedding Planar Graphs Lower Bound). *For any  $\epsilon \in (0, \frac{1}{2})$  and any  $n = \Omega(1/\epsilon^2)$ , there exists an unweighted  $n$ -vertex planar graph  $G(V, E)$  of diameter  $D \leq (1/\epsilon + 2)$  such that for any deterministic dominating embedding  $f : V(G) \rightarrow H$  into a graph  $H$  with additive distortion  $+\epsilon D$ , the treewidth of  $H$  is  $\Omega(1/\epsilon)$ .*

We note that Eisenstat, Klein, and Mathieu [42] implicitly constructed an embedding with additive distortion  $+\epsilon D$  and treewidth  $O(\epsilon^{-1} \cdot \log n)$  (see also [34]). However, in the applications of designing PTASes, the running time is at least exponential in the treewidth, and hence, this embedding only implies inefficient PTASes or QPTASes. The dependency on  $n$  of the treewidth of our embedding in Theorem 3 is exponentially smaller.

Our new embedding result (Theorem 3) leads to an almost linear time PTAS for metric Baker’s problems in planar graphs, thereby giving an affirmative answer to Question 1.

**Theorem 5** (PTAS for Metric Baker Problems in Planar Graphs). *Given an  $n$ -vertex planar graph  $G(V, E, w)$ , two parameters  $\epsilon \in (0, 1)$  and  $\rho > 0$ , and a measure  $\mu : V \rightarrow \mathbb{R}^+$ , one can find in  $2^{\tilde{O}(1/\epsilon^{2+\kappa})} \cdot n^{1+o(1)}$  time for any fixed  $\kappa > 0$ : (1) a  $(1 - \epsilon)\rho$ -independent set  $I$  such that for every  $\rho$ -independent set  $\tilde{I}$ ,  $\mu(I) \geq (1 - \epsilon)\mu(\tilde{I})$  and (2) a  $(1 + \epsilon)\rho$ -dominating set  $S$  such that for every  $\rho$ -dominating set  $\tilde{S}$ ,  $\mu(S) \leq (1 + \epsilon)\mu(\tilde{S})$ .*

We could choose  $\kappa = 0.01$  in Theorem 5 to get a PTAS with running time  $2^{\tilde{O}(1/\epsilon^{2.01})} \cdot n^{1+o(1)}$  for the  $\rho$ -independent set/dominating set problems. The dependency on  $1/\epsilon$  of the running time of our PTAS is much smaller than that of [56].

By using our new embedding in Theorem 3 and some

additional ideas, we significantly improve the running time of the PTAS by Cohen-Addad et al. [34] to almost linear time as asked in Question 2.

**Theorem 6** (PTAS for Bounded-Capacity VRP in Planar Graphs). *There is a randomized polynomial time approximation scheme for the bounded-capacity VRP in planar graphs that runs in time  $O_\epsilon(1) \cdot n^{1+o(1)}$ .*

We note that all results stated for planar graphs so far can be extended to bounded genus graphs with the same bounds using the cutting technique of [34].

In clan embeddings and Ramsey type embeddings, we observe that the construction of the authors [50], in combination with the construction of Cohen-Addad et al. [34], can be seen as providing a reduction from planar graphs with one vortex to  $K_r$ -minor-free graphs. Roughly speaking, the reduction implies that if planar graphs with one vortex have an embedding with treewidth  $t(\epsilon, n)$  and additive distortion  $+\epsilon D$ , then  $K_r$ -minor-free graphs has a clan embedding/Ramsey type embedding with treewidth  $t(O_r(\frac{\delta\epsilon}{\log(n)}), n) + O_r(\log(n))$  and additive distortion  $+\epsilon D$ . The authors [50] used the embedding of planar graphs with one vortex by Cohen-Addad et al. [34], which has treewidth  $t(\epsilon, n) = O(\frac{\log n}{\epsilon})$ , to get a clan embedding/Ramsey type embedding of treewidth  $O_r(\frac{\log n}{(\delta\epsilon)/\log(n)}) + O_r(\log n) = O_r(\frac{\log^2 n}{\delta\epsilon})$ .

There are two fundamental barriers if one wants to improve the treewidth bound of  $O_r(\frac{\log^2 n}{\delta\epsilon})$ : improving the embedding of planar graphs with one vortex and removing the (both multiplicative and additive) loss  $O(\log n)$  in the reduction step. Our embedding for planar graphs with one vortex overcomes the former barrier. Specifically, by plugging in our embedding for planar graphs with one vortex, we obtain treewidth  $O_r(\frac{\log n(\log \log n)^2}{\epsilon\delta})$  in both clan embedding and Ramsey type embedding; the improvement is from quadratic in  $\log(n)$  to nearly linear in  $\log(n)$ . While we are not able to fully answer Question 3 due to the second barrier, we are one step closer to its resolution.

**Theorem 7** (Clan Embeddings of Minor-free Graphs into Low Treewidth Graphs). *Given a  $K_r$ -minor-free  $n$ -vertex graph  $G = (V, E, w)$  of diameter  $D$  and parameters  $\epsilon \in (0, \frac{1}{4})$ ,  $\delta \in (0, 1)$ , there is a distribution  $\mathcal{D}$  over clan embeddings  $(f, \chi)$  with additive distortion  $+\epsilon D$  into graphs of treewidth  $O_r(\frac{\log n(\log \log n)^2}{\epsilon\delta})$  s.t. for every  $v \in V$ ,  $\mathbb{E}[|f(v)|] \leq 1 + \delta$ .*

**Theorem 8** (Ramsey Type Embeddings of Minor-free Graphs into Low Treewidth Graphs). *Given an  $n$ -vertex  $K_r$ -minor-free graph  $G = (V, E, w)$  with diameter  $D$  and parameters  $\epsilon \in (0, \frac{1}{4})$ ,  $\delta \in (0, 1)$ , there is a distribution over dominating embeddings  $g : G \rightarrow H$  into treewidth  $O_r(\frac{\log n(\log \log n)^2}{\epsilon\delta})$  graphs, such that there is a subset  $M \subseteq V$  of vertices for which the following claims hold: noitemsep*

- 1) For every  $u \in V$ ,  $\Pr[u \in M] \geq 1 - \delta$ .
- 2) For every  $u \in M$  and  $v \in V$ ,  $d_H(g(u), g(v)) \leq d_G(u, v) + \epsilon D$ .

Using our new clan and Ramsey embeddings in Theorem 7 and Theorem 8 on top of the algorithm from [50], we improve the running time of the QPTAS for the metric Baker's problems in  $K_r$ -minor-free graphs [50] from  $n^{\tilde{O}(\epsilon^{-2} \log(n))}$  to  $n^{\tilde{O}(\epsilon^{-2}(\log \log(n))^3)}$ .

**Theorem 9** (QPTAS for Metric Baker Problems in Minor-free Graphs). *Given an  $n$ -vertex  $K_r$ -minor-free graph  $G(V, E, w)$ , two parameters  $\epsilon \in (0, 1)$  and  $\rho > 0$ , and a weight function  $\mu : V \rightarrow \mathbb{R}^+$ , one can find in  $n^{\tilde{O}_r(\epsilon^{-2}(\log \log(n))^3)}$  time:*

- (1) a  $(1 - \epsilon)\rho$ -independent set  $I$  such that for every  $\rho$ -independent set  $\tilde{I}$ ,  $\mu(I) \geq (1 - \epsilon)\mu(\tilde{I})$ , and
- (2) a  $(1 + \epsilon)\rho$ -dominating set  $S$  such that for every  $\rho$ -dominating set  $\tilde{S}$ ,  $\mu(S) \leq (1 + \epsilon)\mu(\tilde{S})$ .

## B. Techniques

A key technical component in all the embeddings is an emulator for trees with low treewidth and low hop path between every pair of vertices. For a path  $P$ , denote by  $\text{hop}(P)$  the number of edges in the path  $P$ . For a pair of vertices  $u, v \in V$ , and parameter  $h \in \mathbb{N}$ , denote by

$$d_G^{(h)}(u, v) = \min\{w(P) \mid P \text{ is a } u-v \text{ path s.t. } \text{hop}(P) \leq h\}$$

the weight of a minimum  $u - v$  path with at most  $h$  edges. If no such path exist,  $d_G^{(h)}(u, v) = \infty$ .

Given an edge-weighted tree  $T = (V_T, E_T, w_T)$ , we say that an edge-weighted graph  $K = (V_K, E_K, w_K)$  is an  $h$ -hop emulator for  $T$  if  $V_K = V_T$  and  $d_K^{(h)}(u, v) = d_T(u, v)$  for every pair of vertices  $u, v \in V$ . We show that every tree admits a low-hop emulator with small treewidth.

**Theorem 10.** *Given an edge-weighted  $n$ -vertex tree  $T$ , there is an  $O(\log \log n)$ -hop emulator  $K_T$  for  $T$  that has treewidth  $O(\log \log n)$ . Furthermore,  $K_T$  can be constructed in  $O(n \cdot \log \log n)$  time.*

Constructing a low hop emulator for trees is a basic problem that was studied by various authors [5], [22], [30], [32], [51], [78], [83], [86]. The goal of these constructions is to minimize the number of edges of the emulator for a given hop bound, with no consideration to the treewidth of the resulting graph. To the best of our knowledge, the only previously constructed low-hop emulator with bounded treewidth is a folklore recursive construction of a 2-hop emulator with treewidth  $O(\log n)$ . Our emulator seeks to balance between the treewidth and the hop bound, and achieves an exponential reduction in the treewidth (at the expense of an increase in the hop bound). Due to its fundamental nature, we believe that Theorem 10 will find more applications in other contexts.

a) *Embedding of planar graphs:* To avoid notational clutter, we describe our embedding technique for planar graphs instead of addressing planar graphs with one vortex directly (which is the main motivation of our technique). We will use the low hop emulator in Theorem 10 to obtain a simple and efficient construction as claimed in Theorem 3. Our construction not only addresses Question 1, but it is also robust enough to be easily extensible to planar graphs with

one vortex, which, as discussed above, is the fundamental barrier towards an embedding of  $K_r$ -minor-free graphs with low treewidth.

Given a planar graph  $G = (V, E, w)$  with diameter  $D$ , we construct a recursive decomposition  $\Phi$  of  $G$  using shortest path separators [71], [85]. Specifically, we pick a global root vertex  $r \in V$ .  $\Phi$  has a tree structure of depth  $O(\log n)$ , and each node  $\alpha \in \Phi$  is associated with a cluster of  $G$ , the “border” of which consists of  $O(1)$  shortest paths rooted at  $r$ . This set of border shortest paths is denoted by  $\mathcal{Q}_\alpha$ . For each node  $\alpha$ , all the vertices not in  $\mathcal{Q}_\alpha$  are called *internal vertices*, and denoted  $I_\alpha$ . Each leaf node  $\alpha \in \Phi$  contains only a constant number of internal vertices.

Following the standard techniques [34], [42], for each path  $Q \in \mathcal{Q}_\alpha$  associated with a node  $\alpha \in \Phi$ , one would place a set of equally-spaced vertices of size  $O(\frac{\log(n)}{\epsilon})$ , called *portals*, such that the distance between any two nearby portals is  $O(\epsilon D / \log(n))$ . Since  $|\mathcal{Q}_\alpha| = O(1)$ , each node of  $\Phi$  has  $O(\log(n)/\epsilon)$  portals. We form a tree embedding of *treewidth*  $O(\log(n)/\epsilon)$  from portals of  $\Phi$  by considering every edge  $\{\alpha, \beta\}$  in  $\Phi$  and adding all pairwise edges between portals of  $\alpha$  and  $\beta$ . For the distortion argument, we consider any shortest path  $P$  of  $G$  from  $u$  to  $v$ . Let  $\Phi[\alpha_u, \alpha_v]$  be the path in  $\Phi$  between a node  $\alpha_u$  containing  $u$  and  $\alpha_v$  containing  $v$ . As  $\Phi$  has depth  $O(\log n)$ ,  $\Phi[\alpha_u, \alpha_v]$  has  $O(n)$  nodes. Thus,  $P$  intersects at most  $O(\log n)$  shortest paths on the boundary of  $O(\log n)$  nodes in  $\Phi[\alpha_u, \alpha_v]$ . For any shortest path  $Q \in \mathcal{Q}_\alpha$  that intersects  $P$  for  $\alpha \in \Phi[\alpha_u, \alpha_v]$ , we route  $P$  through the portal closest to the intersection vertex of  $P$  and  $Q$ , incurring an additive distortion of  $O(\epsilon D / \log(n))$ . This means that the total distortion due to routing through portals is  $O(\log n) \cdot O(\epsilon D / \log(n)) = O(\epsilon)D$ .

Our key idea is to use [Theorem 10](#) to construct an  $O(\log \log n)$ -hop emulator  $K_\Phi$  for  $\Phi$  along with a tree decomposition  $\mathcal{T}$  for  $K_\Phi$  of width  $\text{tw}(\mathcal{T}) = O(\log \log n)$ . For each path  $Q \in \mathcal{Q}_\alpha$  in each node  $\alpha \in \Phi$ , we now only place  $O(\frac{\log \log(n)}{\epsilon})$  portals such that the distance between any two nearby portals is  $O(\epsilon D / \log \log(n))$ . For each edge  $\{\alpha, \beta\}$  in  $K_\Phi$ , as opposed to only consider edges in  $\Phi$  in the standard technique, we add all pairwise edges between portals of  $\alpha$  and  $\beta$  to obtain the host graph  $H$ . By replacing each node  $\alpha \in \Phi$  in each bag of  $\mathcal{T}$  with the portals of  $\alpha$ , we obtain a tree decomposition  $\mathcal{X}_H$  of  $H$ . Since  $\mathcal{T}$  has width  $O(\log \log n)$  and each node has  $O(\log \log(n)/\epsilon)$  portals, the treewidth of  $\mathcal{X}_H$  is  $O(\log \log(n)) \cdot O(\log \log(n)/\epsilon) = O((\log \log n)^2/\epsilon)$ .

For the distortion argument, we crucially use the property that  $K_\Phi$  has hop diameter  $O(\log \log(n))$ . Consider any shortest path  $P$  of  $G$ , whose endpoints are internal vertices of two leaf nodes, say  $\alpha_1$  and  $\alpha_2$ , of  $\Phi$ . There is a (shortest) path consisting of  $k = O(\log n \log n)$  nodes from  $\alpha_1$  to  $\alpha_2$  in  $K_\Phi$ . This means that it suffices to consider  $k$  shortest paths  $\{Q_1, \dots, Q_k\}$  associated with these nodes intersecting  $P$ . By the construction, there are edges in  $H$  between portals of  $Q_i$  and  $Q_{i+1}$  for every  $1 \leq i \leq k - 1$ . Thus, by routing an intersection vertex of  $P \cap Q_i$  to the nearest portal of  $Q_i$  for every  $i \in [k]$ , we obtain a path  $P_H$  in  $H$  between the endpoints

of  $P$ . As the portal distance is  $O(\epsilon D / \log \log(n))$ , each routing step incurs  $O(\epsilon D / \log \log(n))$  additive distortion. Since the number of paths is  $k = O(\log \log n)$ , the total distortion is  $k \cdot O(\epsilon D / \log \log(n)) = O(\epsilon)D$ , as desired.

Conceptually, one can think of our shortcutting technique as an effective way to reduce the recursion depth  $O(\log n)$  by balanced separators to  $O(\log \log(n))$ . Since using balanced separators is a fundamental technique in designing planar graphs algorithms, we believe that our idea or its variants could be used to improve other algorithms where the depth  $O(\log n)$  of the recursion tree is the main barrier.

*b) Extension to  $K_r$ -minor-free graphs:* Since planar graphs with one vortex have balanced shortest path separators [34], the embedding algorithm for planar graphs described above is readily applicable to these graphs; the treewidth of the embedding is  $O((\log \log n)^2/\epsilon)$ . If we plug this embedding into the framework of Cohen-Addad et al. [34], we obtain a stochastic embedding of  $K_r$ -minor-free graphs into  $O_r((\log \log n)^2 \epsilon^{-2}) + O_r(\log n)$  treewidth graphs. The additive  $O_r(\log n)$  term is due to a reduction step involving *clique-sums* (Lemma 19 in [34]).

Our idea to remove the  $+O_r(\log n)$  additive term using Robertson-Seymour decomposition is as follows. By Robertson-Seymour decomposition,  $G$  can be decomposed into a tree  $\mathbb{T}$  of nearly embeddable graphs, called *pieces*, that are glued together via *adhesions* of constant size. Our new embedding of planar graphs with one vortex implies that nearly embeddable graphs have embeddings into treewidth  $O((\log \log n)^2/\epsilon)$ . To glue these embeddings together, we first root the tree  $\mathbb{T}$ . Then for each piece  $G_i$ , we simply add the adhesion  $J_0$  between  $G_i$  and its parent bag  $G_0$  to every bag in the tree decomposition of the host graph  $H_i$  in the embedding of  $G_i$ ; this increases the treewidth of  $G_0$  by only  $+O(h(r))$ , and the overall treewidth of the final embedding is also increased only by  $+O(h(r))$ . For the stretch, let  $Q_{uv}$  be a shortest path between two vertices  $u$  and  $v$ . We show by induction that for every vertex  $u \in G_a$  and  $v \in G_b$  for any two pieces  $G_a, G_b$  such that  $G_i$  is their lowest common ancestor, there is a vertex  $x \in Q_{uv} \cap G_a$  and  $y \in Q_{uv} \cap G_b$ , the distances between  $u$  and  $x$ , and between  $v$  and  $y$  are preserved *exactly*. Thus, the distortion for  $d_G(u, v)$  is just  $+\epsilon D$ . Since we apply this construction to *every* piece of  $\mathbb{T}$ ,  $+\epsilon D$  is also the distortion of the final embedding.

*c) Bounded-capacity VRP in planar graphs:* For bounded-capacity VRP in planar graphs, the main ingredient is a rooted-stochastic embedding, where given a root  $r$  (the depot in the VRP problem), there is a distribution over embeddings such that every pair of vertices  $u, v$ , have an additive distortion of  $\epsilon \cdot (d_G(r, u) + d_G(r, v))$  in expectation. Efficient construction of such embedding is the main bottleneck in constructing near-linear time PTAS for the bounded-capacity VRP. The stochastic embedding was introduced by Becker et al. [19]. The main idea of their embedding is to randomly slice graphs into vertex-disjoint bands, construct an additive embedding into a bounded treewidth graph for each band, and finally combine the resulting embeddings. There are two steps in

the algorithm of Becker et al. [19] that incur running time  $\Omega(n^2)$ . First, the algorithm uses the additive embedding of Fox-Epstein et al. [56], which has a large (but polynomial) running time. We improve this step by using the embedding in Theorem 3. Second, for each band  $B$ , they construct a new graph  $G_B$  by taking the union of the shortest paths of all pairs of vertices in  $B$ . Since these shortest paths may contain vertices not in  $B$ , the size of  $G_B$  could be  $\Omega(n)$ ; as a result, the total size of all subgraphs constructed from the bands is  $\Omega(n^2)$ . We improve this step by showing that it suffices to embed the graph induced by the band  $G[B]$  plus one more vertex, and hence the total number of vertices of all subgraphs is  $O(n)$ . The key idea in our proof is to observe that if a vertex falls far from the boundary of the band it belongs to, it has small additive distortion w.r.t. all other vertices. We then carefully choose the parameters so that each vertex is successful in this aspect with probability  $1 - \epsilon$ . Interestingly, we obtain here a strong Ramsey-type guarantee: each vertex  $v$ , with probability  $1 - \epsilon$ , has a small additive distortion w.r.t. all other vertices (instead of a small expected distortion w.r.t. each specific vertex).

### C. Related Work

Different types of embeddings were studied for planar and minor free graphs.  $K_r$ -minor-free graphs embed into  $\ell_p$  space with multiplicative distortion  $O_r(\log^{\min\{\frac{1}{2}, \frac{1}{p}\}} n)$  [2], [3], [47], [72], [81], in particular, they embed into  $\ell_\infty$  of dimension  $O_r(\log^2 n)$  with a constant multiplicative distortion. They also admit spanners with multiplicative distortion  $1 + \epsilon$  and  $\tilde{O}_r(\epsilon^{-3})$  lightness [23]. On the other hand, there are other graph families that embed well into bounded treewidth graphs. Talwar [84] showed that graphs with doubling dimension  $d$  and aspect ratio  $\Phi$  stochastically embed into graphs with treewidth  $\epsilon^{-O(d \log d)} \cdot \log^d \Phi$  and expected distortion  $1 + \epsilon$ . Similar embeddings are known for graphs with highway dimension  $h$  [45] (into treewidth  $(\log \Phi)^{-O(\log^2 \frac{h}{\epsilon})}$  graphs), and graphs with correlation dimension  $k$  [31] (into treewidth  $\tilde{O}_{k,\epsilon}(\sqrt{n})$  graphs).

Ramsey type embeddings of general graphs into ultrametrics and trees were extensively studied [1], [15], [16], [21], [25], [50], [76], [77]. Clan embeddings were also studied for trees [50]. Clan embeddings are somewhat similar to the previously introduced multi-embeddings [14], [26] in that each vertex is mapped to a subset of vertices. However, multi-embeddings lack the notion of chief, which is crucial in all our applications. See also [59]. Finally, clan and Ramsey type embeddings were also studied in the context of hop-constrained metric embeddings [49], [60].

### D. Note on the Full Version

Due to lack of space, we only present the proof of Theorem 10 and Theorem 3 in this extended abstract. Readers are referred to the full version [46] for proofs of all other results.

## II. PRELIMINARIES

$\tilde{O}$  notation hides poly-logarithmic factors, that is  $\tilde{O}(g) = O(g) \cdot \text{polylog}(g)$ , while  $O_r$  notation hides factors in  $r$ , e.g.

$O_r(m) = O(m) \cdot f(r)$  for some function  $f$  of  $r$ . All logarithms are at base 2 (unless specified otherwise).

We consider connected undirected graphs  $G = (V, E)$  with edge weights  $w_G : E \rightarrow \mathbb{R}_{\geq 0}$ . A graph is called unweighted if all its edges have unit weight. Additionally, we denote  $G$ 's vertex set and edge set by  $V(G)$  and  $E(G)$ , respectively. Often, we will abuse notation and write  $G$  instead of  $V(G)$ . Let  $T(V_T, E_T)$  be a tree. For any two vertices  $u, v \in V_T$ , we denote by  $T[u, v]$  the unique path between  $u$  and  $v$  in  $T$ .  $d_G$  denotes the shortest path metric in  $G$ , i.e.,  $d_G(u, v)$  is the shortest distance between  $u$  to  $v$  in  $G$ . When the graph is clear from the context, we might use  $d$  to refer to  $d_G$ .  $G[S]$  denotes the induced subgraph by  $S$ .

A metric embedding is a function  $f : X \rightarrow Y$  between the points of two metric spaces  $(X, d_X)$  and  $(Y, d_Y)$ . A metric embedding  $f$  is said to be *dominating* if for every pair of points  $x, y \in X$ , it holds that  $d_X(x, y) \leq d_Y(f(x), f(y))$ . We say the a metric embedding  $f$  has additive distortion  $+\Delta$  if  $f$  is dominating, and for every  $x, y \in X$ ,  $d_Y(f(x), f(y)) \leq d_X(x, y) + \Delta$ . Stochastic embedding is a distribution  $\mathcal{D}$  over dominating embeddings  $f : V(G) \rightarrow V(H_f)$ . Stochastic embedding  $\mathcal{D}$  has expected additive distortion  $+\Delta$ , if for every pair of points  $x, y \in X$ , it holds that  $\mathbb{E}_{f \sim \mathcal{D}}[d_{H_f}(f(x), f(y))] \leq d_G(x, y) + \Delta$ .

## III. EMULATORS FOR TREES

In this section, we focus on proving Theorem 10. Our proof uses the following lemma, which states that in every tree, we can remove a small number of vertices, such that every remaining connected component has at most two boundary vertices. This lemma was used, sometimes implicitly, in the literature [7], [54], [79]. We include a proof for completeness.

**Lemma 1.** *Given any parameter  $\ell \in \mathbb{N}$  and an  $n$ -vertex tree  $T$ , then in  $O(n)$  time we can find a subset  $X$  of at most  $\frac{2}{\ell+1}n - 1 = O(\frac{n}{\ell})$  vertices such that every connected component  $C$  of  $T \setminus X$  is of size  $|C| \leq \ell$ , and  $C$  has at most two outgoing edges towards  $X$ .*

*Proof.* We begin the proof of Lemma 1 with the following easy claim whose proof can be found in the full version of the paper.

**Claim 1.** *Given an  $n$ -vertex tree  $T$ , in  $O(n)$  time we can find a set  $A$  of at most  $\frac{n}{\ell+1}$  vertices such that every connected component in  $T \setminus A$  contains at most  $\ell$  vertices.*

We first apply Claim 1 to obtain a set  $A$  of size  $\frac{n}{\ell+1}$  in  $O(n)$  time such that every connected component of  $T \setminus A$  has at most  $\ell$  vertices. The set  $X$  contains  $A$  and the least common ancestor of every pair of vertices in  $A$ .  $X$  is called the *least common ancestor closure* of  $A$  (see [54]). Given  $A$ ,  $X$  can be found in  $O(n)$  time by traversing  $T$  in post order. Clearly, every connected component  $C$  of  $T \setminus X$  has size at most  $\ell$ . Note that if  $C$  has outgoing edges towards  $x, y \in X$ , then necessarily  $x$  is an ancestor of  $y$ , or vice versa (as otherwise their least common ancestor will be in  $C$ , but this is a contradiction as  $C \cap X = \emptyset$ ). As it is impossible that

$C$  will have three outgoing edges towards  $x, y, z \in X$  where  $x$  is an ancestor of  $y$ , and  $y$  is an ancestor of  $z$ , we conclude that  $C$  has at most two outgoing edges towards  $X$ .

Finally, we argue by induction on  $n$  that  $|X| \leq 2|A| - 1$ ; the lemma will then follow as  $2|A| - 1 \leq \frac{2}{\ell+1}n - 1$ . The base of the induction when  $|A| = 1$  is clear. Consider a tree  $T$  rooted at a vertex  $r$  with children  $v_1, \dots, v_s$  whose subtrees are denoted by  $T_1, \dots, T_s$ , respectively. First, assume that  $r \notin A$ . If all the vertices of  $A$  belong to a single subtree  $T_i$ , then  $|X| \leq 2|A| - 1$  by the induction hypothesis on  $T_i$ . Otherwise, suppose w.l.o.g. that  $A$  intersects  $T_1, \dots, T_{s'}$  where  $s' \geq 2$ . Then  $X$  will contain the root  $r$ , along with some vertices in each subtree  $T_i$  for  $i \in [1, s']$ . Using the induction hypothesis we conclude that:

$$|X| \leq 1 + \sum_{i=1}^{s'} (2|T_i \cap A| - 1) = 2|A| + 1 - s' \leq 2|A| - 1.$$

The case where  $r \in X$  is similar.  $\square$

We are now ready to prove [Theorem 10](#).

*Proof of Theorem 10.* Fix  $c = \frac{2}{\log \frac{3}{2}}$ . We will prove by induction on  $n$ , that every  $n$ -vertex tree  $T$  for  $n \geq 2$  has a  $(1 + c \cdot \log \log n)$ -hop emulator  $K_T$ , where  $K_T$  has treewidth at most  $1 + c \cdot \log \log n$ . The base case is when  $n = 1$ ; we simply return the original tree  $K_T = T$  (a singleton emulator with treewidth 0). For the general case, using [Lemma 1](#) with parameter  $\ell = \sqrt{2n} - 1$ ,<sup>2</sup> we obtain a set  $X$  of at most  $\frac{2}{\ell+1}n - 1 = \ell$  vertices, such that  $T \setminus X$  has a set of connected components  $\mathcal{C}$ , where every  $C \in \mathcal{C}$  is of size  $|C| \leq \ell$ , and has at most two outgoing edges towards  $X$ . Note that for  $n \geq 1$ ,  $|X| \leq \ell = \sqrt{2n} - 1 < n^{\frac{2}{3}}$ . For each such connected component  $C$ , let  $T_C = T[C]$  be the induced subtree, and  $K_C$  be the emulator constructed using the induction hypothesis, with the tree decomposition  $\mathcal{T}_C$ . Next, we create a new tree  $T_X$  with  $X$  as a vertex set as follows. For every connected component  $C$  with two outgoing edges towards two vertices  $u, v \in X$ , we add an edge  $\{u, v\}$  to  $T_X$  of weight  $d_T(u, v)$ . One can easily verify that  $T_X$  is indeed a tree (in fact it is a minor of  $T$ ), and furthermore, for every two vertices  $x, y \in X$ ,  $d_{T_X}(x, y) = d_T(x, y)$ . We recursively construct an emulator  $K_X$  along with a tree decomposition  $\mathcal{T}_X$  for  $T_X$ .

We construct the emulator  $K_T$  for  $T$  as follows:  $K_T$  contains the emulator  $K_X$ , the union of all the emulators  $\cup_{C \in \mathcal{C}} K_C$ , and in addition, for every  $C \in \mathcal{C}$  with outgoing edges towards  $\{u, v\}$  (we allow  $u = v$ ), and for every vertex  $z \in C$ , we add edges  $\{z, u\}, \{z, v\}$  to  $K_T$ . (The weight of every edge added to  $K_T$  is the distance between their endpoints in  $T$ .)

First, we argue that  $K_T$  has low hop. Consider a pair of vertices  $u, v$  in  $T$ , and suppose that  $u \in C^u$  and  $v \in C^v$  where  $C^u$  and  $C^v$  are in  $\mathcal{C}$ . (The cases where either  $u$  or  $v$  is in  $X$  follows by the same argument.) If  $C^v = C^u$ , then by the induction hypothesis on  $K_{C^u}$ ,  $d_{K_T}^{(1+c \cdot \log \log n)}(u, v) \leq$

$d_{K_{C^u}}^{(1+c \cdot \log \log |C^u|)}(u, v) = d_{T_{C^u}}(u, v) = d_T(u, v)$ . Else, the unique  $u - v$  path  $P$  in  $T$  must go through  $X$ . Let  $x_u, x_v \in P \cap X$  be the closest vertices to  $u$  and  $v$ , respectively. Since  $K_T$  contains the edges  $\{u, x_u\}, \{v, x_v\}$ , by the induction hypothesis on  $K_X$ , we have:

$$d_{K_T}^{(3+c \cdot \log \log |X|)}(u, v) \leq d_T(u, x_u) + d_{K_X}^{(1+c \cdot \log \log |X|)}(x_u, x_v) + d_T(x_v, v) = d_T(u, v).$$

Observe that  $3 + c \cdot \log \log |X| \leq 3 + c \cdot \log \log n^{\frac{2}{3}} = 3 + c \cdot \log \frac{2}{3} + c \cdot \log \log n = 1 + c \cdot \log \log n$ . This means that  $d_{K_T}^{(1+c \cdot \log \log n)}(u, v) \leq d_{K_T}^{(3+c \cdot \log \log |X|)}(u, v) \leq d_T(u, v)$ , as desired.

Next, we argue that  $K_T$  has treewidth at most  $1 + c \cdot \log \log n$ . We define a tree-decomposition  $\mathcal{T}$  as follows: For every cluster  $C \in \mathcal{C}$  with outgoing edges towards  $\{u, v\}$  (we allow  $u = v$ ), let  $\mathcal{T}_C$  be the tree decomposition obtained by the induction hypothesis on  $T[C]$ . Let  $\mathcal{T}'_C$  be the tree decomposition obtained from  $\mathcal{T}_C$  by adding the vertices  $u, v$  to every bag. Consider the tree decomposition  $\mathcal{T}_X$  of  $K_X$ . Since  $K_X$  contains the edge  $\{u, v\}$ , there is a bag  $B_{u,v} \in \mathcal{T}_X$  containing both  $u$  and  $v$ . We add an edge in  $\mathcal{T}$  between  $B_{u,v}$  and an arbitrary bag in  $\mathcal{T}_C$ . It follows directly from the construction that  $\mathcal{T}$  is a valid tree decomposition of  $K_T$ . Furthermore, the width of the decomposition is bounded by

$$\max\{1 + c \cdot \log \log |X|, 3 + c \cdot \log \log \ell\} \leq 1 + c \cdot \log \log n,$$

as required.

Finally, we bound the running time. Note that by [Lemma 1](#), we can find  $X$  in  $O(n)$  time. Constructing the set of connected components  $\mathcal{C}$  and the tree  $T_X$  takes  $O(n)$  time. To find the weights for edges of  $T_X$ , we use an exact distance oracle for trees with construction time  $O(n)$  and query time  $O(1)$ . It is folklore that such a distance oracle can be constructed by a simple reduction to the least common ancestor (LCA) data structure [\[20\]](#) (see also [\[63\]](#), [\[82\]](#)). Constructing  $K_T$  from  $K_X$  and  $\{K_C\}_{C \in \mathcal{C}}$  takes  $O(n)$  time as well. As the depth of the recursion is  $O(\log \log n)$ , and the total running time of each level is  $O(n)$ , the overall running time is  $O(n \log \log n)$ .  $\square$

#### IV. EMBEDDING PLANAR GRAPHS

In this section, we focus on proving [Theorem 3](#). Throughout, we fix a planar drawing of  $G$ . W.l.o.g, we assume that  $G$  is triangulated; otherwise, we can triangulate  $G$  in linear time and set the weight of the new edges to be  $+\infty$ . Let  $r$  be an (arbitrary) vertex of  $G$  and  $T_r$  be a shortest path tree rooted at  $r$ . We say that a shortest path  $Q$  in  $G$  is an  $r$ -path if  $Q$  is a path in  $T_r$  with  $r$  as an endpoint.

We start by defining  $\eta$ -rooted shortest path decomposition.

**Definition 1** ( $\eta$ -RSPD). *An  $\eta$ -rooted shortest path decomposition with root  $r$ , denoted by  $\Phi$  of a graph is a binary tree with the following properties: noitemsep*

- (P1.)  $\Phi$  has height  $O(\log n)$  and  $O(n)$  nodes.
- (P2.) Each node  $\alpha \in \Phi$  is associated with a subgraph  $X_\alpha$  of  $G$ , called a piece, such that:

<sup>2</sup>For the sake of simplicity, we will ignore integral issues.

- (a)  $X_\alpha$  contains at most  $\eta$  shortest paths rooted at  $r$ , called boundary paths. Denote by  $\mathcal{Q}_\alpha$  be the set of all boundary paths. We will abuse notation and denote by  $\mathcal{Q}_\alpha$  the union of all the vertices in all the boundary paths. Every vertex in  $X_\alpha$  which does not lie on a boundary path, is called an internal vertex.
- (b) If  $\alpha$  is the root of  $\Phi$ , then  $X_\alpha = G$ ; if  $\alpha$  is a leaf of  $\Phi$ , then  $X_\alpha$  has at most  $\eta$  internal vertices. Otherwise,  $\alpha$  is an internal node with exactly two children  $\beta_1, \beta_2$ . It holds that  $X_\alpha = X_{\beta_1} \cup X_{\beta_2}$  and  $X_{\beta_1} \cap X_{\beta_2}$  are the boundary paths shared by  $X_{\beta_1}$  and  $X_{\beta_2}$ .
- (c) For any vertex  $u \in V(X_\alpha)$  and  $v \in V(G) \setminus V(X_\alpha)$ , (any) path between  $u$  and  $v$  must intersect some vertex that lies on a boundary path of  $X_\alpha$ .

In planar graphs, an  $\eta$ -RSPD is also known as a *recursive decomposition* with shortest path separators, see, e.g., [8], [67], [71]. Abraham et al. [2] (see also [47]) defined a related notion of shortest path decomposition (SPD). There are several differences between RSPD and SPD : SPD the height of the tree is a general parameter  $k$ ; the number of boundary paths in each node is equal to its depth (distance to the root); the tree is not necessarily binary; and the leaves contain no internal node. SPD was used to construct multiplicative embeddings into  $\ell_1$  [2], [47] and scattering partitions [48]. The following observation follows directly from the definition.

For any two nodes  $\alpha, \beta \in \Phi$ , let  $\Phi[\alpha, \beta]$  be the subpath between  $\alpha$  and  $\beta$  in  $\Phi$ . A crucial property of of an RSPD in our construction is the *separation property*: for any path  $P_{uv}$  between two vertices  $u \in \mathcal{Q}_\alpha$  and  $v \in \mathcal{Q}_\beta$  of two given nodes  $\alpha$  and  $\beta$ , for any node  $\lambda$  on the path between  $\alpha$  and  $\beta$  in the tree  $\Phi$ ,  $P_{uv} \cap \mathcal{Q}_\lambda \neq \emptyset$ .

An explicit representation of  $\Phi$  could take  $\Omega(n^2)$  bits and hence is not computable in nearly linear time. However,  $\Phi$  has a *compact representation* that only takes  $O(n\eta)$  words of space. Specifically, we store at each node  $\alpha \in \Phi$  at most  $\eta$  vertices  $B_\alpha = \{v_1, v_2, \dots, v_{\eta'}\}$  such that  $\{T_r[r, v_i]\}_{i=1}^{\eta'}$  is the set of  $r$ -paths  $\mathcal{Q}_\alpha$ . If  $\alpha$  is a leaf node of  $\Phi$ , then  $\alpha$  is associated with an extra set of  $O(1)$  vertices, denoted by  $I_\alpha$ , that are internal vertices of  $X_\alpha$ . Using shortest path separators, Thorup [85] showed that one can compute a compact representation of an  $O(1)$ -RSPD of  $G$  in time  $O(n \log n)$ . (See Section 2.5 in [85] for details; the RSPD is called *frame separator decomposition* of  $G(V, E, w)$  in Thorup's paper.)

We now present an embedding of  $G(V, E, w)$  into a low-treewidth graph in nearly linear time. We refer readers to Section I-B for an overview of the proof. We will construct a one-to-many embedding  $f : V \rightarrow 2^{V(H)}$  with additive distortion  $+O(\epsilon) \cdot D$ ; we can recover distortion  $\epsilon D$  by scaling  $\epsilon$ . We are required to guarantee that for any two copies of a single vertex  $v_1, v_2 \in f(v)$ ,  $d_H(v_1, v_2) = O(\epsilon) \cdot D$ . In the end, we can transform  $f$  to a one-to-one embedding  $f'$  by picking an arbitrary copy  $v' \in f(v)$  and set  $f'(v) = v'$  for each vertex  $v \in V(G)$ .

Recall that  $T_r$  is a shortest path tree rooted at  $r$  of

$G(V, E, w)$ . Since  $G(V, E, w)$  has diameter  $D$ ,  $T_r$  has radius  $D$ . Let  $\delta = \frac{\epsilon D}{\log \log n}$ . We first define a set of vertices  $N_r$  called  $\delta$ -portals of  $T_r$  as follows: initially,  $N_r$  only contains  $r$ ; we then visit every vertex of  $T_r$  in the depth-first order, and we add a vertex  $v$  to  $N_r$  if the nearest ancestor of  $v$  in  $N_r$  is at a distance larger than  $\delta$  from  $v$ . For each  $r$ -path  $Q$  in  $T_r$ , we denote by  $N(Q, \delta) = N_r \cap V(Q)$  the set of  $\delta$ -portals in  $Q$ . Note that for every vertex  $v \in Q$ , there is a  $\delta$ -portal  $u \in N(Q, \delta)$  at a distance at most  $\delta$ . Furthermore, the distance between a pair of consecutive  $\delta$ -portals of  $Q$  is greater than  $\delta$ . As the length of  $Q$  is bounded by  $D$ , we conclude that

$$|N(Q, \delta)| \leq \left\lfloor \frac{D}{\delta} \right\rfloor \leq \frac{\log \log n}{\epsilon}.$$

The construction works as follows: let  $\Phi$  be an  $O(1)$ -RSPD of  $G(V, E, w)$ . For every node  $\alpha \in \Phi$ , let  $Q_1, Q_2, \dots$ , be the  $\eta = O(1)$   $r$ -paths constituting  $\mathcal{Q}_\alpha$ . Denote by  $P_\alpha = \cup_{Q \in \mathcal{Q}_\alpha} N(Q, \delta)$  the union of all  $\delta$ -portals from all the  $r$ -paths on the boundary of  $\alpha$ . Note that  $|P_\alpha| = O(\eta \cdot \frac{\log \log n}{\epsilon}) = O(\frac{\log \log n}{\epsilon})$ . Using Theorem 10, we construct an  $O(\log \log n)$ -hop emulator  $K_\Phi$  for  $\Phi$  with treewidth  $O(\log \log n)$ . We add the following sets of edges to the host graph  $H$ :

- 1) For every node  $\alpha \in \Phi$ , and vertex  $v \in P_\alpha$ , create a copy  $\widetilde{v}_\alpha$ . Denote this set of copies by  $\widetilde{P}_\alpha$ . Add the edge set  $\widetilde{P}_\alpha \times \widetilde{P}_\alpha$  to  $H$  (i.e. a clique on the copies).
- 2) For every edge  $\{\alpha, \beta\} \in K_\Phi$ , add the edge set  $\widetilde{P}_\alpha \times \widetilde{P}_\beta$  to  $H$  (i.e. a bi-clique between the respective copies).
- 3) For every leaf node  $\alpha \in \Phi$ , let  $I_\alpha$  be the set of internal vertices. We add to  $H$  the vertices in  $I_\alpha$  and two edge sets  $I_\alpha \times I_\alpha$  and  $I_\alpha \times \widetilde{P}_\alpha$ .
- 4) Denote by  $I_\Phi = \cup_{\alpha \in \Phi: \alpha \text{ is leaf}} I_\alpha$  the set of all vertices that belong to the interior of leaves of  $\Phi$ . For  $v \notin I_\Phi$ , it necessarily belongs to some  $r$ -path  $Q_v$  on the boundary of some leaf node  $\alpha_v$ . We add  $v$  and the edge set  $v \times \widetilde{P}_{\alpha_v}$  to  $H$ .

In summary, the set of vertices of  $H$  is  $V_H = V_G \cup \cup_{\alpha \in \Phi} \widetilde{P}_\alpha$ . The image  $f(v)$  of every vertex  $v$  consists of  $v$  itself (added to  $H$  in either Step 3 or Step 4), and all the respective copies (added in Step 1, one per each node  $\alpha$  such that  $v \in P_\alpha$ ). The weights of the edges in  $H$  are defined in a natural way: for every edge  $(u', v')$  in  $H$ ,  $w_H(u', v') = d_G(u, v)$  where  $u = f^{-1}(u')$  and  $v = f^{-1}(v')$ . The following lemma bounds the treewidth of  $H$ ; the proof is delayed to the full version [46].

**Lemma 2.**  $H$  has treewidth  $O(\epsilon^{-1}(\log \log n)^2)$ .

We now bound the distortion of the embedding. Specifically, we will show that for every pair of vertices  $u, v$ , and every two copies  $u' \in f(u)$ ,  $v' \in f(v)$ ,  $d_G(u, v) \leq d_H(u', v') \leq d_G(u, v) + O(\epsilon) \cdot D$ . The lower bound follows directly from the way we assign weights to the edges of  $H$ . Indeed, every  $u'-v'$  path in  $H$  corresponds to a  $u-v$  walk in  $G$  of the same weight. Henceforth, we focus on proving the upper bound. First, we need the following lemma, which is the generalization of the separation property.

**Lemma 3.** Let  $\alpha$  and  $\beta$  be two nodes in  $\Phi$ . Let  $P_{uv}$  be any path between two vertices  $u$  and  $v$  in  $G$  such that  $u \in Q_\alpha$  and  $v \in Q_\beta$ . Let  $(\alpha = \lambda_1, \lambda_2, \dots, \lambda_k = \beta)$  be a set of nodes in  $\Phi[\alpha, \beta]$  such that  $\lambda_{i+1} \in \Phi[\lambda_i, \beta]$  for any  $1 \leq i \leq k-1$ . Then, there exists a sequence of vertices  $(u = x_1, x_2, \dots, x_k = v)$  such that  $x_i \in P_{uv} \cap Q_{\lambda_i}$  and  $x_{i+1} \in P_{uv}[x_i, v]$  for any  $1 \leq i \leq k-1$ .

Next, we show that the distortion of portal vertices on the boundaries of nodes in  $\Phi$  is in check.

**Lemma 4.** Let  $\alpha$  and  $\beta$  be two nodes in  $\Phi$ . Let  $\widetilde{u}_\alpha \in \widetilde{P}_\alpha$  and  $\widetilde{v}_\beta \in \widetilde{P}_\beta$  be two vertices in  $H$ , and  $u = f^{-1}(\widetilde{u}_\alpha), v = f^{-1}(\widetilde{v}_\beta)$ . Then,  $d_H(\widetilde{u}_\alpha, \widetilde{v}_\beta) \leq d_G(u, v) + O(\epsilon) \cdot D$

*Proof.* Let  $Q_{uv}$  be a shortest path from  $u$  to  $v$  in  $G$ . Recall that  $K_\Phi$  is an emulator of  $\Phi$  with hop diameter  $O(\log \log n)$ . Let  $P = (\alpha = \lambda_1, \lambda_2, \dots, \lambda_k = \beta)$  be a shortest path from  $\alpha$  to  $\beta$  in  $K_\Phi$  such that  $k = O(\log \log n)$ . Since  $K_\Phi$  preserves distances between nodes of  $\Phi$ , i.e.,  $w_{K_\Phi}(P) = w_\Phi(\Phi[\alpha, \beta])$ , it must be that the nodes on  $P$  constitute a subsequence of nodes on  $\Phi[\alpha, \beta]$ . By Lemma 3, there exists a sequence of vertices  $(u = x_1, x_2, \dots, x_k = v)$  such that  $x_i \in Q_{\lambda_i} \cap Q_{uv}$  for all  $1 \leq i \leq k$  and:

$$d_G(u, v) = \sum_{i=1}^{k-1} d_G(x_i, x_{i+1}) \quad (2)$$

For every  $i \in [k]$ , let  $y_i \in P_{\lambda_i}$  be the  $\delta$ -portal closest to  $x_i$ ; note that  $y_1 = u$  and  $y_k = v$  because  $u$  and  $v$  are  $\delta$ -portals. By the definition of  $\delta$ -portals,  $d_G(x_i, y_i) \leq \delta$ . Thus, by the triangle inequality, it holds that:

$$d_G(y_i, y_{i+1}) \leq d_G(x_i, x_{i+1}) + 2\delta \quad (3)$$

Recall that  $\widetilde{(y_i)}_{\lambda_i}$  be the copy of  $y_i$  created for  $\lambda_i$ . Observe that, since  $(\lambda_i, \lambda_{i+1})$  is an edge in  $K_\Phi$ , by construction, there is an edge between  $\widetilde{(y_i)}_{\lambda_i}$  and  $\widetilde{(y_{i+1})}_{\lambda_{i+1}}$  of weight  $w_H(\widetilde{(y_i)}_{\lambda_i}, \widetilde{(y_{i+1})}_{\lambda_{i+1}}) = d_G(y_i, y_{i+1})$ . Thus, we have:

$$\begin{aligned} d_H(\widetilde{u}_\alpha, \widetilde{v}_\beta) &= d_H(\widetilde{(y_1)}_{\lambda_1}, \widetilde{(y_k)}_{\lambda_k}) \\ &\leq \sum_{i=1}^{k-1} d_H(\widetilde{(y_i)}_{\lambda_i}, \widetilde{(y_{i+1})}_{\lambda_{i+1}}) \\ &= d_G(u, v) + 2(k-1)\delta = d_G(u, v) + O(\epsilon)D, \end{aligned}$$

as claimed.  $\square$

**Lemma 5.** For any  $u, v \in V(G)$  and any  $u' \in f(u), v' \in f(v)$ , it holds that:

$$d_H(u', v') \leq d_G(u, v) + O(\epsilon)D.$$

*Proof.* The proof is by case analysis. If both  $u, v$  are  $\delta$ -portals, then the lemma holds by Lemma 4. Next, suppose that both  $u, v$  are not  $\delta$ -portals. There are two leaf nodes  $\alpha, \beta \in \Phi$ , such that  $u \in X_\alpha$ , and  $v \in X_\beta$ . In particular, there are unique copies of both  $u$  and  $v$ , one for each vertex. If  $\alpha = \beta$ , then  $H$  contains the edge  $(u, v)$  and we are done. Otherwise, let  $P_{uv}$  be a shortest  $u$ - $v$  path in  $G$ . Then it follows from Definition 1 that

there are two vertices  $x_u$  and  $x_v$  in  $P_{uv}$  such that  $x_u \in Q_\alpha$  and  $x_v \in Q_\beta$  (it might be that  $u = x_u$  or  $v = x_v$ ). Furthermore, there are  $\delta$ -portals  $y_u \in P_\alpha$ , and  $y_v \in P_\beta$ , at distances at most  $\delta$  from  $x_u$  and  $x_v$  respectively. By construction in Step 2 and Step 4 (depending on whether  $u, v$  are leaves),  $H$  contains the edges  $(u, \widetilde{(y_u)}_\alpha), (v, \widetilde{(y_v)}_\beta)$ . By Lemma 4, we have:

$$\begin{aligned} d_H(u, v) &\leq d_H(u, \widetilde{(y_u)}_\alpha) + d_H(\widetilde{(y_u)}_\alpha, \widetilde{(y_v)}_\beta) \\ &\quad + d_H(\widetilde{(y_v)}_\beta, v) \leq d_G(u, v) + O(\epsilon) \cdot D. \end{aligned}$$

Finally, the cases where either  $u$  or  $v$  is a  $\delta$ -portal are simpler and can be proved by the same argument.  $\square$

We now compute the runtime of our algorithm. First, we compute a (compact representation of)  $O(1)$ -RSPD  $\Phi$  in  $O(n \log n)$  time. Next, we use Theorem 10 to compute an emulator  $K_\Phi$  in  $O(n \cdot \log \log n)$  time, with  $O(|\Phi| \cdot \log \log |\Phi|) = O(n \cdot \log \log n)$  edges. Next, we compute the set of  $\delta$ -portals. We will use the same shortest path tree  $T_r$  rooted at  $r$  (that can be found in  $O(n)$  time [64]). For each vertex  $v \in T_r$ , we compute and store a set of  $\delta$ -portals  $N(T_r[r, v], \delta)$  of the  $r$ -rooted shortest path  $T_r[r, v]$ ; this can be done in  $O(n \log n \log(n)/\epsilon)$  by a simple tree traversal.

Next, we construct the one-to-many embedding  $f$  into  $H$  with tree-decomposition  $\mathcal{T}$  following to Steps (1)-(4). The construction is straightforward and takes  $O\left(n \cdot \frac{(\log \log n)^3}{\epsilon^2}\right)$  time. The last and most time-consuming step is to assign weights to the edges in  $H$ . Recall that the weight of an edge in  $(u', v') \in H$  where  $u = f^{-1}(u')$  and  $v = f^{-1}(v')$ , is defined to be  $w_H(u', v') = d_G(u, v)$ . Computing the shortest distance between  $u$  and  $v$  in  $G$  takes  $\Omega(n)$  time, and using shortest path computation to find the weight of every edge in  $H$  incurs  $\Omega(n^2)$  time. Our solution is to use approximate distances instead of the exact ones to assign to edges of  $H$ . Specifically, we will use the *approximate distance oracle* of Thorup<sup>3</sup>.

To assign weights in our graph  $H$ , we simply construct a  $(1 + \epsilon)$ -distance oracle  $\mathcal{O}_{G, \epsilon}(u, v)$  for  $G$ , and then for every edge  $(u', v') \in H$  where  $u = f^{-1}(u')$  and  $v = f^{-1}(v')$ , we set  $w_H(u', v') = \mathcal{O}_{G, \epsilon}(u, v)$ . Thus the total time for assigning the weights is  $O\left(\frac{n \log^3(n)}{\epsilon^2}\right) + |E(H)| \cdot O\left(\frac{1}{\epsilon}\right) = O\left(\frac{n \log^3(n)}{\epsilon^2}\right)$ , which is also the overall running time of our algorithm.

#### ACKNOWLEDGMENT

The authors are grateful to Philip Klein and Vincent Cohen-Addad for helpful conversations. We thank anonymous reviewers for helpful comments. In particular, one reviewer pointed out the least common ancestor closure and reference [54] which help us simplify the proof of Theorem 10 considerably.

#### REFERENCES

- [1] Ittai Abraham, Shiri Chechik, Michael Elkin, Arnold Filtser, and Ofer Neiman. Ramsey spanning trees and their applications. *ACM Trans. Algorithms*, 16(2):19:1–19:21, 2020. 7

<sup>3</sup>A similar distance oracle was constructed independently by Klein [71]

- [2] Ittai Abraham, Arnold Filtser, Anupam Gupta, and Ofer Neiman. Metric embedding via shortest path decompositions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018*, pages 952–963, 2018. 7, 9
- [3] Ittai Abraham, Cyril Gavoille, Anupam Gupta, Ofer Neiman, and Kunal Talwar. Cops, robbers, and threatening skeletons: Padded decomposition for minor-free graphs. *SIAM J. Comput.*, 48(3):1120–1145, 2019. preliminary version published in STOC 2014. 7
- [4] A. Adamaszek, A. Czumaj, and A. Lingas. PTAS for  $k$ -tour cover problem on the plane for moderately large values of  $k$ . In *International Symposium on Algorithms and Computation, ISAAC '09*, pages 994–1003, 2009. 2
- [5] N. Alon and B. Schieber. *Optimal preprocessing for answering on-line product queries*. 1987. see here. 5
- [6] Noga Alon, Yossi Matias, and Mario Szegedy. The space complexity of approximating the frequency moments. *J. Comput. Syst. Sci.*, 58(1):137–147, 1999. 1
- [7] S. Alstrup, J. Holm, K. D. Lichtenberg, and M. Thorup. Maintaining information in fully dynamic trees with top trees. *ACM Transactions on Algorithms*, 1(2):243–264, 2005. 7
- [8] Sanjeev Arora, Michelangelo Grigni, David R. Karger, Philip N. Klein, and Andrzej Woloszyn. A polynomial-time approximation scheme for weighted planar graph TSP. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, 25-27 January 1998, San Francisco, California, USA*, pages 33–41, 1998. see here. 4, 9
- [9] Vassilis Athitsos and Stan Sclaroff. Database indexing methods for 3d hand pose estimation. In *Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15-17, 2003, Selected Revised Papers*, pages 288–299, 2003. 1
- [10] B. S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *Journal of the ACM*, 41(1):153–180, 1994. preliminary version published in FOCS 1983. 1, 2
- [11] Yair Bartal. Probabilistic approximations of metric spaces and its algorithmic applications. In *37th Annual Symposium on Foundations of Computer Science, FOCS '96, Burlington, Vermont, USA, 14-16 October, 1996*, pages 184–193, 1996. 1
- [12] Yair Bartal. On approximating arbitrary metrics by tree metrics. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 161–168, 1998. 1
- [13] Yair Bartal. Graph decomposition lemmas and their role in metric embedding methods. In *Algorithms - ESA 2004, 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, pages 89–97, 2004. 1
- [14] Yair Bartal. Advances in metric ramsey theory and its applications. *CoRR*, abs/2104.03484, 2021. 7
- [15] Yair Bartal, Béla Bollobás, and Manor Mendel. Ramsey-type theorems for metric spaces with applications to online problems. *J. Comput. Syst. Sci.*, 72(5):890–921, 2006. Special Issue on FOCS 2001. 7
- [16] Yair Bartal, Nathan Linial, Manor Mendel, and Assaf Naor. On metric ramsey-type phenomena. *Annals of Mathematics*, 162(2):643–709, 2005. 7
- [17] A. Becker. A tight  $4/3$  approximation for capacitated vehicle routing in trees. In 118, editor, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*, pages 3:1–3:15, 2018. 2
- [18] A. Becker, P. N. Klein, and D. Saulpic. A quasi-polynomial-time approximation scheme for vehicle routing on planar and bounded-genus graphs. In *25th Annual European Symposium on Algorithms*, volume 87 of *ESA '17*, pages 12:1–12:15, 2017. 2
- [19] A. Becker, P. N. Klein, and A. Schild. A PTAS for bounded-capacity vehicle routing in planar graphs. In *Algorithms and Data Structures, WADS '19*, pages 99–111, 2019. 2, 3, 6, 7
- [20] M. A. Bender and M. Farach-Colton. The LCA problem revisited. In *Latin American Symposium on Theoretical Informatics, LATIN '00*, pages 88–94, 2000. 8
- [21] Guy E. Blelloch, Yan Gu, and Yihan Sun. A new efficient construction on probabilistic tree embeddings. *CoRR*, abs/1605.04651, 2016. 7
- [22] Hans L. Bodlaender, Gerard Tel, and Nicola Santoro. Trade-offs in non-reversing diameter. *Nord. J. Comput.*, 1(1):111–134, 1994. 5
- [23] G. Borradaile, H. Le, and C. Wulff-Nilsen. Minor-free graphs have light spanners. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science, FOCS '17*, pages 767–778, 2017. 7
- [24] Glencora Borradaile and Hung Le. Optimal dynamic program for  $r$ -domination problems over tree decompositions. In *11th International Symposium on Parameterized and Exact Computation, IPEC 2016, August 24-26, 2016, Aarhus, Denmark*, pages 8:1–8:23, 2016. 2
- [25] J. Bourgain, T. Figiel, and V. Milman. On Hilbertian subsets of finite metric spaces. *Israel J. Math.*, 55(2):147–152, 1986. 7
- [26] John M. Boyer and Wendy J. Myrvold. On the cutting edge: Simplified  $O(n)$  planarity by edge addition. *J. Graph Algorithms Appl.*, 8(3):241–273, 2004. 7
- [27] Sébastien Bubeck, Michael B. Cohen, Yin Tat Lee, James R. Lee, and Aleksander Madry.  $k$ -server via multiscale entropic regularization. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 3–16, 2018. 1
- [28] Douglas E. Carroll and Ashish Goel. Lower bounds for embedding into distributions over excluded minor graph families. In *Algorithms - ESA 2004, 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, pages 146–156, 2004. 1
- [29] Amit Chakrabarti, Alexander Jaffe, James R. Lee, and Justin Vincent. Embeddings of topological graphs: Lossy invariants, linearization, and 2-sums. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 761–770, 2008. 1
- [30] T.-H. H. Chan and A. Gupta. Small hop-diameter sparse spanners for doubling metrics. *Discrete & Computational Geometry*, 41(1):28–44, 2008. 5
- [31] T.-H. Hubert Chan and Anupam Gupta. Approximating TSP on metrics with bounded global growth. *SIAM J. Comput.*, 41(3):587–617, 2012. preliminary version published in SODA 2008. 7
- [32] B. Chazelle. Computing on a free tree via complexity-preserving mappings. *Algorithmica*, 2(1-4):337–361, 1987. 5
- [33] V. Cohen-Addad, É. Colin de Verdière, P. N. Klein, C. Mathieu, and D. Meierfrankenfeld. Approximating connectivity domination in weighted bounded-genus graphs. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 584–597. ACM, 2016. 4
- [34] Vincent Cohen-Addad, Arnold Filtser, Philip N. Klein, and Hung Le. On light spanners, low-treewidth embeddings and efficient traversing in minor-free graphs. In *61th Annual IEEE Symposium on Foundations of Computer Science, FOCS*, pages 589–600, 2020. 2, 3, 4, 5, 6
- [35] George B Dantzig and John H Ramser. The truck dispatching problem. *Management science*, 6(1):80–91, 1959. 2
- [36] A. Das and C. Mathieu. A quasipolynomial time approximation scheme for euclidean capacitated vehicle routing. *Algorithmica*, 73(1):115–142, 2015. 2
- [37] E. D. Demaine and M. Hajiaghayi. Bidimensionality: New connections between FPT algorithms and PTASs. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA'05*, pages 590–601, 2005. 1
- [38] E. D. Demaine, M. Hajiaghayi, and K. Kawarabayashi. Algorithmic graph minor theory: Decomposition, approximation, and coloring. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 637–646, 2005. 1, 2
- [39] E. D. Demaine, M. Hajiaghayi, and B. Mohar. Approximation algorithms via contraction decomposition. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '07*, pages 278–287, 2007. 1
- [40] Erik D. Demaine, Fedor V. Fomin, Mohammad Taghi Hajiaghayi, and Dimitrios M. Thilikos. Fixed-parameter algorithms for  $(k, r)$ -center in planar graphs and map graphs. *ACM Trans. Algorithms*, 1(1):33–47, 2005. preliminary version published in ICALP 2003. 2
- [41] Erik D. Demaine, MohammadTaghi Hajiaghayi, and Ken-ichi Kawarabayashi. Contraction decomposition in  $h$ -minor-free graphs and algorithmic applications. In *Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC 2011, San Jose, CA, USA, 6-8 June 2011*, pages 441–450, 2011. 1
- [42] David Eisenstat, Philip N. Klein, and Claire Mathieu. Approximating  $k$ -center in planar graphs. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 617–627, 2014. 4, 6
- [43] D. Eppstein. Diameter and treewidth in minor-closed graph families. *Algorithmica*, 27(3):275–291, 2000. 1
- [44] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. Syst.*

- Sci.*, 69(3):485–497, November 2004. preliminary version published in STOC 2003. 1
- [45] Andreas Emil Feldmann, Wai Shing Fung, Jochen Könnemann, and Ian Post. A  $(1+\epsilon)$ -embedding of low highway dimension graphs into bounded treewidth graphs. *SIAM J. Comput.*, 47(4):1667–1704, 2018. preliminary version published in ICALP 2015. 7
- [46] A. Filtser and H. Le. Low treewidth embeddings of planar and minor-free metrics. *Arxiv Preprint*, 2022. <https://arxiv.org/abs/2203.15627>. 7, 9
- [47] Arnold Filtser. A face cover perspective to  $\ell_1$  embeddings of planar graphs. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 1945–1954. SIAM, 2020. 7, 9
- [48] Arnold Filtser. Scattering and sparse partitions, and their applications. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of *LIPIcs*, pages 47:1–47:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 9
- [49] Arnold Filtser. Hop-constrained metric embeddings and their applications. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 492–503. IEEE, 2021. 7
- [50] Arnold Filtser and Hung Le. Clan embeddings into trees, and low treewidth graphs. In *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, 2021*, pages 342–355. ACM, 2021. 2, 3, 4, 5, 7
- [51] Arnold Filtser and Hung Le. Reliable spanners: Locality-sensitive orderings strike back. In Anupam Gupta, editor, *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, June 20-24, 2022*. ACM, 2022. to appear, see [full version](#). 5
- [52] M. Fisher. Chapter 1 vehicle routing. In *Handbooks in Operations Research and Management Science*, pages 1–33. Elsevier, 1995. 2
- [53] F. V. Fomin, D. Lokshtanov, V. Raman, and S. Saurabh. Bidimensionality and eptas. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '11*, page 748–759, 2011. 1
- [54] F. V. Fomin, D. Lokshtanov, S. Saurabh, and M. Zehavi. *Kernelization*. Cambridge University Press, December 2018. 7, 10
- [55] K. Fox, P. N. Klein, and S. Mozes. A polynomial-time bicriteria approximation scheme for planar bisection. In *Proceedings of the 47th annual ACM symposium on Theory of Computing*, 2015. 4
- [56] E. Fox-Epstein, P. N. Klein, and A. Schild. Embedding planar graphs into low-treewidth graphs with applications to efficient approximation schemes for metric problems. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '19*, page 1069–1088, 2019. 1, 2, 3, 4, 7
- [57] Lee-Ad Gottlieb, Aryeh Kontorovich, and Robert Krauthgamer. Efficient regression in metric spaces via approximate lipschitz extension. *IEEE Trans. Inf. Theory*, 63(8):4838–4849, 2017. preliminary version published in SIMBAD 2013. 1
- [58] Michelangelo Grigni, Elias Koutsoupias, and Christos H. Papadimitriou. An approximation scheme for planar graph TSP. In *36th Annual Symposium on Foundations of Computer Science, Milwaukee, Wisconsin, USA, 23-25 October 1995*, pages 640–645, 1995. 4
- [59] Bernhard Haeupler, D. Ellis Hershkowitz, and Goran Zuzic. Deterministic tree embeddings with copies for algorithms against adaptive adversaries. *CoRR*, abs/2102.05168, 2021. 7
- [60] Bernhard Haeupler, D. Ellis Hershkowitz, and Goran Zuzic. Tree embeddings for hop-constrained network design. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 356–369. ACM, 2021. 7
- [61] Mordecai Haimovich and Alexander H. G. Rinnooy Kan. Bounds and heuristics for capacitated routing problems. *Math. Oper. Res.*, 10(4):527–542, 1985. 2
- [62] Eran Halperin, Jeremy Buhler, Richard M. Karp, Robert Krauthgamer, and B. Westover. Detecting protein sequence conservation via metric embeddings. *Bioinformatics*, 19(suppl 1):i122–i129, 07 2003. 1
- [63] Dov Harel and Robert Endre Tarjan. Fast algorithms for finding nearest common ancestors. *SIAM J. Comput.*, 13(2):338–355, 1984. 8
- [64] M. R. Henzinger, P. Klein, S. Rao, and S. Subramanian. Faster shortest-path algorithms for planar graphs. *Journal of Computer and System Sciences*, 55(1):3–23, 1997. 10
- [65] A. Jayaprakash and M. R. Salavatipour. Approximation schemes for capacitated vehicle routing on graphs of bounded treewidth, bounded doubling, or highway dimension. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '22*, pages 877–893, 2022. 2
- [66] Ioannis Katsikarelis, Michael Lampis, and Vangelis Th. Paschos. Structurally parameterized d-scattered set. *Discrete Applied Mathematics*, 2020. preliminary version published in WG 2018. 2
- [67] K. Kawarabayashi, P. N. Klein, and C. Sommer. Linear-space approximate distance oracles for planar, bounded-genus and minor-free graphs. In *The 38th International Colloquium on Automata, Languages and Programming, ICALP '11*, pages 135–146, 2011. 9
- [68] K. Kawarabayashi, Y. Kobayashi, and B. Reed. The disjoint paths problem in quadratic time. *Journal of Combinatorial Theory, Series B*, 102(2):424–435, 2012. 3
- [69] Maleq Khan, Fabian Kuhn, Dahlia Malkhi, Gopal Pandurangan, and Kunal Talwar. Efficient distributed approximation algorithms via probabilistic tree embeddings. *Distributed Comput.*, 25(3):189–205, 2012. preliminary version published in PODC 2008. 1
- [70] P. N. Klein. A linear-time approximation scheme for planar weighted TSP. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, FOCS '05*, pages 647–657, 2005. 1
- [71] Philip N. Klein. Preprocessing an undirected planar network to enable fast approximate distance queries. In *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 6-8, 2002, San Francisco, CA, USA*, pages 820–827, 2002. 6, 9, 10
- [72] Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured descent: a new embedding method for finite metrics. *Geometric and Functional Analysis*, 15(4):839–858, 2005. preliminary version published in FOCS 2004. 7
- [73] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Comb.*, 15(2):215–245, 1995. preliminary version published in FOCS 1994. 1
- [74] Dániel Marx and Michal Pilipczuk. Optimal parameterized algorithms for planar facility location problems using voronoi diagrams. In *Algorithms - ESA 2015 - 23rd Annual European Symposium, Patras, Greece, September 14-16, 2015, Proceedings*, pages 865–877, 2015. 2
- [75] C. Mathieu and H. Zhou. A PTAS for capacitated vehicle routing on trees, 2021. 2
- [76] Manor Mendel and Assaf Naor. Ramsey partitions and proximity data structures. *Journal of the European Mathematical Society*, 9(2):253–275, 2007. preliminary version published in FOCS 2006. 7
- [77] Assaf Naor and Terence Tao. Scale-oblivious metric fragmentation and the nonlinear dvoretzky theorem. *Israel Journal of Mathematics*, 192(1):489–504, 2012. 7
- [78] Giri Narasimhan and Michiel H. M. Smid. *Geometric spanner networks*. Cambridge University Press, 2007. 5
- [79] J. Nederlof, M. Pilipczuk, C. M. F. Swennenhuis, and K. Węgrzycki. Isolation Schemes for Problems on Decomposable Graphs. In *39th International Symposium on Theoretical Aspects of Computer Science (STACS 2022)*, Leibniz International Proceedings in Informatics (LIPIcs), pages 50:1–50:20, 2022. 7
- [80] C. H. Papadimitriou and M. Yannakakis. The Traveling Salesman Problem with distances one and two. *Mathematics of Operations Research*, 18(1):1–11, 1993. 2
- [81] Satish Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the Fifteenth Annual Symposium on Computational Geometry, Miami Beach, Florida, USA, June 13-16, 1999*, pages 300–306, 1999. 7
- [82] B. Schieber and U. Vishkin. On finding lowest common ancestors: Simplification and parallelization. *SIAM Journal on Computing*, 17(6):1253–1262, 1988. 8
- [83] Shay Solomon. An optimal-time construction of sparse euclidean spanners with tiny diameter. In Dana Randall, editor, *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 820–839. SIAM, 2011. 5
- [84] Kunal Talwar. Bypassing the embedding: algorithms for low dimensional metrics. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, 2004*, pages 281–290, 2004. 7
- [85] M. Thorup. Compact oracles for reachability and approximate distances in planar digraphs. *Journal of the ACM*, 51(6):993–1024, 2004. 6, 9
- [86] Mikkel Thorup. Parallel shortcutting of rooted trees. *Journal of Algorithms*, 23(1):139–159, 1997. 5